

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:49:11 ; Search time 44 Seconds  
(without alignments)  
4322.307 Million cell updates/sec

Title: US-09-497-822C-19

Perfect score: 4912

Sequence: 1 MEVQLGLGRVYPPRPSTYR.....SVQPKILSGKVKPIVHTQ 923

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	%	Query Length	DB ID	Description
1	4777.5	97.3	906	4	Q9UN21	Q9un21 homo sapien
2	4341.5	88.4	896	6	Q9GKL7	Q9gkl7 sus scrofa
3	4298	87.5	895	6	Q9GKN9	Q9gkn9 sus scrofa
4	3822.5	77.8	730	4	Q13771	Q13771 homo sapien
5	2873	58.5	544	4	Q9BZG7	Q9bzg7 homo sapien
6	2857	58.2	542	4	Q9BZG6	Q9bzg6 homo sapien
7	2846.5	57.9	539	4	Q9NUA2	Q9nu22 homo sapien
8	2802.5	57.1	531	4	Q9BZG5	Q9bzg5 homo sapien
9	2414.5	49.2	790	13	P700048	P70048 xenopus lae
10	1676	34.1	344	13	Q91445	Q91445 serinus can
11	1639.5	33.4	853	13	Q93245	Q93245 oncorhynch
12	1632.5	33.2	839	13	Q9QFV2	Q9qfv2 carassius a
13	1604	32.7	303	6	Q97684	Q97684 ovis aries
14	1604	32.7	769	13	Q93497	Q93497 pagrus majo
15	1600	32.6	797	13	Q9PWG5	Q9pws5 anguilla ja
16	1556	31.7	854	13	Q93244	Q93244 oncorhynch

17	1552	31.6	763	13	Q9QFV7	Q9qfv7 haplochromi
18	1538.5	31.3	763	13	Q8UWB7	Q8uwb7 oreochromis
19	1532	31.2	848	13	Q9YGV9	Q9ygv9 anguilla ja
20	1472.5	30.0	563	13	Q9DDJ4	Q9ddj4 halichoeres
21	1270	25.9	939	6	Q9GLW0	Q9glw0 canis famil
22	1209	24.6	688	13	Q8UWB8	Q8uwb8 oreochromis
23	1180.5	24.0	692	13	Q9W6F4	Q9w6f4 haplochromi
24	1179	24.0	732	13	Q9DDU9	Q9ddu9 xenopus lae
25	1177.5	24.0	703	13	Q8UVY3	Q8uvy3 xenopus lae
26	1157.5	23.6	583	13	Q9DEV4	Q9dev4 xenopus lae
27	1150	23.4	232	13	Q9IA30	Q9ia30 anolis caro
28	1125	22.9	710	13	Q9IBD5	Q9ibd5 anguilla ja
29	1112	22.6	438	13	Q9OZM7	Q9ozm7 petromyzon
30	1090.5	22.2	794	11	Q8R463	Q8r463 rattus norv
31	1082	22.0	793	11	Q8R5J0	Q8r5j0 rattus norv
32	1064.5	21.7	360	13	Q42274	Q42274 crocodylus
33	1055.5	21.5	689	13	Q8QGX5	Q8qgx5 anguilla ja
34	1055	21.5	982	6	Q9N0W8	Q9n0w8 saimiri sci
35	1016.5	20.7	401	13	Q9OZM6	Q9ozm6 petromyzon
36	998	20.3	831	4	Q8TDS3	Q8tds3 homo sapien
37	978	19.9	232	6	Q9BE98	Q9be98 sus scrofa
38	966	19.7	348	13	Q91425	Q91425 cnemidophor
39	892	18.2	359	13	Q9IAC6	Q9iac6 oncorhynch
40	871	17.7	164	13	Q9OZC2	Q9ozc2 calotes ver
41	837	17.0	284	13	Q9OY00	Q9oy00 petromyzon
42	816	16.6	298	6	Q8S547	Q8s547 ovis aries
43	730	14.9	149	6	Q8SQ22	Q8sq22 capra hircu
44	725.5	14.8	200	13	Q918F5	Q918f5 plimephales
45	715	14.6	166	13	Q91698	Q91698 xenopus lae

## ALIGNMENTS

RESULT 1

Q9UN21 PRELIMINARY; PRT; 906 AA.

ID Q9UN21; AC Q9UN21; DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Androgen receptor.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI\_TaxID=9606;

[1]

SEQUENCE FROM N.A.

RC TISSUE-BREAST CARCINOMA;

RA Jin C.H., Urcan-Bisel M.S., Schrader W.T.;

RT "Androgen receptor sequences in human mammary carcinoma MDA-MB-453 cells."

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.

DR EMBL; AF162704; AAD45921.1; .

DR HSSP; P06536; 1GDC.

DR InterPro; IPR001103; Andrgn\_receptor.

DR InterPro; IPR000536; Hormone\_rec\_lig.

DR InterPro; IPR001628; Znf\_C4steroid.

DR Pfam; PF02166; Androgen\_recep; 1.

DR Pfam; PF00104; hormone\_rec; 1.

DR Pfam; PF00105; zf-C4; 1.

DR PRINTS; PR00047; STROIDFINGER.

DR ProDom; PD000035; Znf\_C4steroid; 1.

DR SMART; SM00430; HOL1; 1.

DR SMART; SM00399; Znf\_C4; 1.

DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.

KW DNA-binding; Nuclear protein; Receptor; Transcription regulation; Zinc-finger.

SQ SEQUENCE 906 AA; 97884 MW; 0A1FA8802B2EDDAF CRC64;

Query Match 97.3%; Score 4777.5; DB 4; Length 906;



```

Db 401 LASLHGGGAPGPGSGSPSATSSSSWHTLFTAESQLYGPC-----GGGG 445
Qy 473 GGGGAGAVAGYGYTRPPQGLAGQESDFTAPDVWYPGMVSRVYPSPTCVKSEMGPWMD 532
Db 446 GSAGEAGAVAGYGYTRPPQGLAGQEGDLAPDIWIPGVVSRVYPSPCVKSEMGPWME 505
Qy 533 SYSGPYGDMRLTARDHVLPIIDYFPQKTCICGDEASGCHYGALTCGSKCVFFKRAAE 592
Db 506 SYSGPYGDMRLTARDHVLPIIDYFPQKTCICGDEASGCHYGALTCGSKCVFFKRAAE 565
Qy 593 GKQYLCAASRNDCTIDFRKNCPSCLRKCIEYAGMTLGARKLKLGKLNKLQEGEASST 652
Db 566 GKQYLCAASRNDCTIDFRKNCPSCLRKCIEYAGMTLGARKLKLGKLNKLQEGEASST 625
Qy 653 TSPTTEETQKLTVSHIEGYECQPIFLNVLEAIEPVGVCAGHDNNQDPSFAALLSSNELG 712
Db 626 TSPTTEPAQKLTVSHIEGYECQPIFLNVLEAIEPVGVCAGHDNNQDPSFAALLSSNELG 685
Qy 713 ERQLVHVYVWAKALPGFRNLHVDQMAVIOYSWMGLMVFAMGWSFTNVNSRMLYFAPDL 772
Db 686 ERQLVHVYVWAKALPGFRNLHVDQMAVIOYSWMGLMVFAMGWSFTNVNSRMLYFAPDL 745
Qy 773 VFNEYRMHKSRYMSQCVMRHLSOEFGWLQITPQEFCLMKALLFSIIPVDGLKNQKFFD 832
Db 746 VFNEYRMHKSRYMSQCVMRHLSOEFGWLQITPQEFCLMKALLFSIIPVDGLKNQKFFD 805
Qy 833 ELRMNYIKELDRIIACKRNKPTSCSRRFYQLTLLDSVQPIARELHQTFFDLIKSHMVS 892
Db 806 ELRMNYIKELDRIIACKRNKPTSCSRRFYQLTLLDSVQPIARELHQTFFDLIKSHMVS 865
Qy 893 VDFPEMAEIIISVQPKILSGKVKPIYFHTQ 923
Db 866 VDFPEMAEIIISVQPKILSGKVKPIYFHTQ 896

RESULT 3
Q9GKN9 PRELIMINARY; PRT; 895 AA.
ID Q9GKN9;
AC Q9GKN9;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Androgen receptor AR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_taxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20539123; PubMed=11086548;
RA Trakooljul N., Ponsuksilli S., Schellander K., Wimmers K.;
RT "A highly polymorphic repetitive polypyrimidine/polypurine (CCTTT)n
RT sequence in the 5' untranslated sequence of the porcine androgen
RT receptor gene.";
RL Anim. Genet. 31:288-289(2000).
CC -!- SUBCELLULAR LOCATION: NUCLEAR [BY SIMILARITY].
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; AF161717; AAG40566.1; -.
DR HSSP; P06536; 1GDC.
DR InterPro; IPR001103; Andrgn_receptor.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001628; ZnF_C4steroid.
DR Pfam; PF02166; Androgen_recep. 1.
DR Pfam; PF00104; hormone_rec. 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; ZnF_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; ZnF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 895 AA; 97156 MW; 923C2FDD1F7E4779 CRC64;

```

```

Query Match 87.5%; Score 4298; DB 6; Length 895.
Best Local Similarity 87.7%; Pred No. 5, 9e-289;
Matches 816; Conservative 19; Mismatches 53; Indels 42; Gaps 4;

Qy 1 MEVQLGLGRVYPRPSTKYRGAFQNLFSVREVIVQNGRHPPEAASAPPGASLLLLQQ 60
Db 1 MEVQLGLGRVYWPWPSKTRFGAFQNLFSVREVIVQNGRHPPEAASAPPGARL----- 54
Qy 61 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQ--GEDGSPQAHRRGPTGYLVLEDEQOP 118
Db 55 -----QQQQQQQQQQQQQQQQQQQQQQQQQQQQQSPQVQSGRGTGYLVLEDEQOP 100
Qy 119 SQPSALECHPERGCVPEPAAVAASKGLPQQLPAPPDEDDSAAPSTLSLLGTPPGGLSS 178
Db 101 SQQSAPECHPEGCTPEPAAVAASKGLQOQPAPPDEDDSAAPSTLSLLGTPPGGLSS 160
Qy 179 CSAADLKDIILSEASTMOLL-----QQQQQAVSEGGSSSGRAREASGAPTSSKNYLGST 233
Db 161 CSTDLKDIILSEAGTMOLLQQQQQQQQQQQQQQQQQAVSEGGSSSGRAREATGAPISCKDSYLGSSST 220
Qy 234 ISDNAKELCKAVSVSMGLGVEALEHLSPEQLRGDCMVAPLLGVPPAVRPTCAPLAECK 293
Db 221 ISDNAKELCKAVSVSMGLGVEALEHLSPEQLRGDCMVAPLLGVPPAVRPTCAPLAECK 280
Qy 294 GSLLDSAGCKSTEDTAEYSPFKGGYTKGLESGSLGSCGSAAGSSGTLPLSTLSLYKSG 353
Db 281 GSLLDGPCKSNEETAEYSPFKAGYTKGLDSESLGSSGSGEGSGTLELPSALSLYKSG 340
Qy 354 ALDEAAAYOSRDYNNPPLALAGPPPPPPHPPHARIKLENPLDYGSAWAAAAQRYGDL 413
Db 341 ALDDVAAYPSRDYNNPPLALARPPPPPPHPPHARIKLENPLDYGSAWAAAAQRYGDL 400
Qy 414 ASLHGAGAAGPGSGSPSAASSSWHTLFTAEGQLYGPCGGGGGGGGGGGGGGGG 473
Db 401 ASLHGAGAGSGSGSPSATSSSSWHTLFTAESQLYGPC-----GGGGG 445
Qy 474 GGGEAGAVAGYGYTRPPQGLAGQESDFTAPDVWYPGMVSRVYPSPTCVKSEMGPWMD 533
Db 446 SAGEAGAVAGYGYTRPPQGLAGQEGDLAPDIWIPGVVSRVYPSPCVKSEMGPWME 505
Qy 534 YSGPYGDMRLTARDHVLPIIDYFPQKTCICGDEASGCHYGALTCGSKCVFFKRAAG 593
Db 506 YSGPYGDMRLTARDHVLPIIDYFPQKTCICGDEASGCHYGALTCGSKCVFFKRAAG 565
Qy 594 KQYLCASRNDCTIDFRKNCPSCLRKCIEYAGMTLGARKLKLGKLNKLQEGEASST 653
Db 566 KQYLCASRNDCTIDFRKNCPSCLRKCIEYAGMTLGARKLKLGKLNKLQEGEASST 625
Qy 654 SPTEETQKLTVSHIEGYECQPIFLNVLEAIEPVGVCAGHDNNQDPSFAALLSSNELGE 713
Db 626 SPTEEPAQKLTVSHIEGYECQPIFLNVLEAIEPVGVCAGHDNNQDPSFAALSSNELGE 685
Qy 714 RQLVHVYVWAKALPGFRNLHVDQMAVIOYSWMGLMVFAMGWSFTNVNSRMLYFAPDL 773
Db 686 RQLVHVYVWAKALPGFRNLHVDQMAVIOYSWMGLMVFAMGWSFTNVNSRMLYFAPDL 745
Qy 774 FNEYRMHKSRYMSQCVMRHLSOEFGWLQITPQEFCLMKALLFSIIPVDGLKNQKFFD 833
Db 746 FNEYRMHKSRYMSQCVMRHLSOEFGWLQITPQEFCLMKALLFSIIPVDGLKNQKFFD 805
Qy 834 LRMYIKELDRIIACKRNKPTSCSRRFYQLTLLDSVQPIARELHQTFFDLIKSHMVS 893
Db 806 LRMYIKELDRIIACKRNKPTSCSRRFYQLTLLDSVQPIARELHQTFFDLIKSHMVS 865
Qy 894 DPEPMAEIIISVQPKILSGKVKPIYFHTQ 923
Db 866 DPEPMAEIIISVQPKILSGKVKPIYFHTQ 895

RESULT 4
Q13771 PRELIMINARY; PRT; 730 AA.
ID Q13771

```

```
AC Q13771;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Androgen receptor.
GN AR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=90258935; PubMed=2342476;
RA Govindan M.V.;
RT "Specific region in hormone binding domain is essential for hormone
RT binding and trans-activation by human androgen receptor.";
RL Mol. Endocrinol. 4:417-427(1990).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; M73069; AAA51735.1; -.
DR HSSP; P06536; IGDC.
DR InterPro; IPR001103; Andrgn_receptor.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF02166; Androgen_rec; 1.
DR Pfam; PF00104; hormone_rec; 1.
DR PRINTS; PF00105; zf-C4; 1.
DR PROSITE; PS00047; STROIDFINGER.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 730 AA; 79150 MW; 918847C3B41B80C9 CRC64;

Query Match 77.8%; Score 3822.5; DB 4; Length 730;
Best Local Similarity 97.7%; Pred. No. 3.7e-256;
Matches 721; Conservative 0; Mismatches 2; Indels 15; Gaps 2;

QY 193 MQLLQQQQQAVSEGGSSGRAREASGAPTSKDNLYGGTSTSDNAKELCKAVSVSMGLG 252
DB 1 MQLLQQQQQAVSEGGSSGRAREASGAPTSKDNLYGGTSTSDNAKELCKAVSVSMGLG 60
QY 253 VEALHLSPEQLRGDCMYAPLLGVPAVRPTCAPLAECKSLDDSDSAGKSTEDTAEYS 312
DB 61 VEALHLSPEQLRGDCMYAPLLGVPAVRPTCAPLAECKSLDDSDSAGKSTEDTAEYS 120
QY 313 PFKGGYTKLGESLGCSSAAGSSGTLPLSTLSLYKSGALDAAAYQSRDYNFPLA 372
DB 121 PFKGGYTKLGESLGCSSAAGSSGTLPLSTLSLYKSGALDAAAYQSRDYNFPLA 180
QY 373 LAGPPPPPPPHPHARIKLENPLDYGSAWAAAACRYGDLASLHGAAGPGSGPSAA 432
DB 181 LAGPPPPPPPHPHARIKLENPLDYGSAWAAAACRYGDLASLHGAAGPGSGPSAA 240
QY 433 ASSSWHTLFTAERGQLYGPCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 492
DB 241 ASSSWHTLFTAERGQLYGPCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 292
QY 493 LAQESDFTAPDWWYVPGWVSRVPYPTCVKSEMGPMWDSYSGPYGDMRLTARDHVL 552
DB 293 LAQESDFTAPDWWYVPGWVSRVPYPTCVKSEMGPMWDSYSGPYGDMRLTARDHVL 352
QY 553 IDYFFPQKTCICGDEASGCHYGALTCGSKVFFKRAEAGKOKYLCASRNDCTDKFRR 612
DB 353 IDYFFPQKTCICGDEASGCHYGALTCGSKVFFKRAEAGKOKYLCASRNDCTDKFRR 412
QY 613 KNCPSRLKRCYAGMTLGARKLKLGNLKLQEGEASGTTSTPTETTKLTVSHIEGYE 672
DB 413 KNCPSRLKRCYAGMTLGARKLKLGNLKLQEGEASGTTSTPTETTKLTVSHIEGYE 472
```

```
QY 673 COPIFLNVLEAIEPGVWCAGHDNNQPDFAALLSSLNELGERQLVHVVKWAKALPGFRL 732
DB 473 COPIFLNVLEAIEPGVWCAGHDNNQPDFAALLSSLNELGERQLVHVVKWAKALPGFRL 532
QY 733 HVDQMAVIOYQSWGLAVFANGWRSFTNVNSRMILYFAPDLVFNFNEYRMHKSMTYSCQVMR 792
DB 533 HVDQMAVIOYQSWGLAVFANGWRSFTNVNSRMILYFAPDLVFNFNEYRMHKSMTYSCQVMR 592
QY 793 HLSQEFGLQITPQEFELCMKA-----LLLSFIIPVDGLKNQKFFDELRMNYIKELDRI 845
DB 593 HLSQEFGLQITPQEFELCMKAMLIFFFLLSFIIPVDGLKNQKFFDELRMNYIKELDRI 652
QY 846 IACKKKNPTSCSRFYQTLKLLDSVQPIARELHQFTFDLLIKSHMVSVDPEMMAEIIISV 905
DB 653 IACKKKNPTSCSRFYQTLKLLDSVQPIARELHQFTFDLLIKSHMVSVDPEMMAEIIISV 712
QY 906 QVPKILSKGVPIYFHTQ 923
DB 713 QVPKILSKGVPIYFHTQ 730

RESULT 5
Q9BZG7 PRELIMINARY; PRT; 544 AA.
ID Q9BZG7 AC Q9BZG7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Androgen receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Megklara A., Diamandis E.P.;
RT "poly-Q and poly-G repeats in the androgen receptor of the breast
RT cancer cell line ZR-75-1.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF321914; AAK09423.1; -.
DR InterPro; IPR001103; Andrgn_receptor.
DR Pfam; PF02166; Androgen_recep; 1.
DR Receptor.
KW NON_TER 544 544
SQ SEQUENCE 544 AA; 56013 MW; C97133EB6C922E74 CRC64;

Query Match 58.5%; Score 2873; DB 4; Length 544;
Best Local Similarity 99.6%; Pred. No. 1.3e-190;
Matches 542; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MEVOLGLGRVYPRPPSKTYRGAFONLFQSVREVIONPGRHPEAASAAAPPGASILL--Q 58
DB 1 MEVOLGLGRVYPRPPSKTYRGAFONLFQSVREVIONPGRHPEAASAAAPPGASILLQQQ 60
QY 59 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 118
DB 61 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 120
QY 119 SOPSALCEHPERCVCVPEPGAANAASKGLFQOLPAPDEDDSAAPSTLSLGGTFPGGLSS 178
DB 121 SOPSALCEHPERCVCVPEPGAANAASKGLFQOLPAPDEDDSAAPSTLSLGGTFPGGLSS 180
QY 179 CSADLKIDILSEATMQLLQQQQQEAIVSEGGSSGRAREASGAPTSKDNLYGGTSTISDNA 238
DB 181 CSADLKIDILSEATMQLLQQQQQEAIVSEGGSSGRAREASGAPTSKDNLYGGTSTISDNA 240
QY 239 KELCKAVSVSMGLGVEALEHLSPEQLRGDCMYAPLLGVPAVRPTCAPLAECKSLDD 298
DB 241 KELCKAVSVSMGLGVEALEHLSPEQLRGDCMYAPLLGVPAVRPTCAPLAECKSLDD 300
QY 299 DSAGKSTEDTAEYSPFKGGYTKLGESLGCSSAAGSSGTLPLSTLSLYKSGALDEA 358
DB 301 DSAGKSTEDTAEYSPFKGGYTKLGESLGCSSAAGSSGTLPLSTLSLYKSGALDEA 360
```

```
QY 359 AAYQSRDYNNFPLALAGPPPPPPPPHARIKLENPLDYGSAWAAAAAQCRCYGDILASLHG 418
|||||
Db 361 AAYQSRDYNNFPLALAGPPPPPPPPHARIKLENPLDYGSAWAAAAAQCRCYGDILASLHG 420
|||||
QY 419 AGAAGPGSGSPSAASSSWHTLFTAEGLYPCGGGGGGGGGGGGGGGGGGGGGA 478
|||||
Db 421 AGAAGPGSGSPSAASSSWHTLFTAEGLYPCGGGGGGGGGGGGGGGGGGGGGA 480
|||||
QY 479 GAVAPGYTRPPQGLAQESDFTAPDVWYPGGMVSRVPYPTCVKSEMPWMDSYSGPY 538
|||||
Db 481 GAVAPGYTRPPQGLAQESDFTAPDVWYPGGMVSRVPYPTCVKSEMPWMDSYSGPY 540
|||||
QY 539 GDMR 542
|||||
Db 541 GDMR 544

RESULT 6
Q9BZG6 ID Q9BZG6 PRELIMINARY; PRT; 542 AA.
AC Q9BZG6;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Androgen receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Magklara A., Diamandis E.P.;
RT "Poly-Q and poly-G repeats in the androgen receptor of the breast
RT cancer cell line BT-474."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF321916; AAK09425.1;
DR InterPro; IPR001103; Andrgn_receptor.
DR Pfam; PF02166; Androgen_recip; 1.
KW Receptor.
FT NON_TER 542 542
SQ SEQUENCE 542 AA; 55628 MW; C363EF841CAAF7739 CRC64;

Query Match 58.2%; Score 2857; DB 4; Length 542;
Best Local Similarity 99.6%; Pred. No. 1.6e-189;
Matches 541; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 MEVOLGLGRVYPRPSKTYRGAFONLFQSVREVIONPGRHPEAASAPPGASLLILQQ 59
|||||
Db 1 MEVOLGLGRVYPRPSKTYRGAFONLFQSVREVIONPGRHPEAASAPPGASLLILQQ 60
|||||
QY 60 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 119
|||||
Db 61 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 120
|||||
QY 120 QPSALECHPERGCVPEPGAASAKGLPQOLPAPDEDDSAAPSTLSLLGPTFPGLSSC 179
|||||
Db 121 QPSALECHPERGCVPEPGAASAKGLPQOLPAPDEDDSAAPSTLSLLGPTFPGLSSC 180
|||||
QY 180 SADLKDILSEASTWQLLQOQQQQAASVSGSSGRAREASGAPTSKDNLYLGGTSTISDNK 239
|||||
Db 181 SADLKDILSEASTWQLLQOQQQQAASVSGSSGRAREASGAPTSKDNLYLGGTSTISDNK 240
|||||
QY 240 ELCKAVSVSMGLGVEALEHLSFGELRGDCMYPALGVPVAVRTPCAPLAECCKGSLDD 299
|||||
Db 241 ELCKAVSVSMGLGVEALEHLSFGELRGDCMYPALGVPVAVRTPCAPLAECCKGSLDD 300
|||||
QY 300 SAGKSTEDTAESYSPFKGGYTKGLESGLCGSGSAAAGSSGTLEPSTLSLYKSGALDEAA 359
|||||
Db 301 SAGKSTEDTAESYSPFKGGYTKGLESGLCGSGSAAAGSSGTLEPSTLSLYKSGALDEAA 360
|||||
QY 360 AYQSRDYNNFPLALAGPPPPPPPPHARIKLENPLDYGSAWAAAAAQCRCYGDILASLHG 419
|||||
```

```
Db 361 AYQSRDYNNFPLALAGPPPPPPPPHARIKLENPLDYGSAWAAAAAQCRCYGDILASLHG 420
|||||
QY 420 GAAGPGSGSPSAASSSWHTLFTAEGLYPCGGGGGGGGGGGGGGGGGGGGGA 479
|||||
Db 421 GAAGPGSGSPSAASSSWHTLFTAEGLYPCGGGGGGGGGGGGGGGGGGGGGA 479
|||||
QY 480 AVAPYGYTRPPQGLAQESDFTAPDVWYPGGMVSRVPYPTCVKSEMPWMDSYSGPY 539
|||||
Db 480 AVAPYGYTRPPQGLAQESDFTAPDVWYPGGMVSRVPYPTCVKSEMPWMDSYSGPY 539
|||||
QY 540 DMR 542
|||||
Db 540 DMR 542

RESULT 7
Q9NUA2 ID Q9NUA2 PRELIMINARY; PRT; 539 AA.
AC Q9NUA2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Androgen receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Magklara A., Diamandis E.P.;
RT "Poly-Q and poly-G repeats in the androgen receptor of the breast
RT cancer cell line T-47D."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF321915; AAK09424.1;
DR InterPro; IPR001103; Andrgn_receptor.
DR Pfam; PF02166; Androgen_recip; 1.
KW Receptor.
FT NON_TER 539 539
SQ SEQUENCE 539 AA; 55444 MW; AB493953B89D869F CRC64;

Query Match 57.9%; Score 2846.5; DB 4; Length 539;
Best Local Similarity 99.4%; Pred. No. 8.7e-189;
Matches 539; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 1 MEVOLGLGRVYPRPSKTYRGAFONLFQSVREVIONPGRHPEAASAPPGASLLILQQ 60
|||||
Db 1 MEVOLGLGRVYPRPSKTYRGAFONLFQSVREVIONPGRHPEAASAPPGASLLILQQ 58
|||||
QY 61 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 120
|||||
Db 59 QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ 118
|||||
QY 121 POSALECHPERGCVPEPGAASAKGLPQOLPAPDEDDSAAPSTLSLLGPTFPGLSSC 180
|||||
Db 119 POSALECHPERGCVPEPGAASAKGLPQOLPAPDEDDSAAPSTLSLLGPTFPGLSSC 178
|||||
QY 181 ADLKDILSEASTWQLLQOQQQQAASVSGSSGRAREASGAPTSKDNLYLGGTSTISDNK 240
|||||
Db 179 ADLKDILSEASTWQLLQOQQQQAASVSGSSGRAREASGAPTSKDNLYLGGTSTISDNK 238
|||||
QY 241 LCKAVSVSMGLGVEALEHLSFGELRGDCMYPALGVPVAVRTPCAPLAECCKGSLDD 300
|||||
Db 239 LCKAVSVSMGLGVEALEHLSFGELRGDCMYPALGVPVAVRTPCAPLAECCKGSLDD 298
|||||
QY 301 SAGKSTEDTAESYSPFKGGYTKGLESGLCGSGSAAAGSSGTLEPSTLSLYKSGALDEAA 360
|||||
Db 299 SAGKSTEDTAESYSPFKGGYTKGLESGLCGSGSAAAGSSGTLEPSTLSLYKSGALDEAA 358
|||||
QY 361 AYQSRDYNNFPLALAGPPPPPPPPHARIKLENPLDYGSAWAAAAAQCRCYGDILASLHG 420
|||||
Db 359 AYQSRDYNNFPLALAGPPPPPPPPHARIKLENPLDYGSAWAAAAAQCRCYGDILASLHG 418
|||||
QY 421 AAGPGSGSPSAASSSWHTLFTAEGLYPCGGGGGGGGGGGGGGGGGGGGGA 480
|||||
```

```
Db 419 AAGCGSSPAAASSWHTLFTAEGLYGPC-CGGGGGGGGGGGGGGGGGGGGEAGA 477
Qy 481 VAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPTCVKSEMGPMWDSYSGPYGD 540
Db 478 VAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPTCVKSEMGPMWDSYSGPYGD 537
Qy 541 MR 542
Db 538 MR 539

RESULT 8
ID Q9BZG5 PRELIMINARY; PRT; 531 AA.
AC Q9BZG5;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Androgen receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Magklara A., Diamandis E.P.;
RT "Poly-Q and poly-G repeats in the androgen receptor of the breast
RT cancer cell line MCF-7."
RL EMBL: AF321917; AAK09426.1; to the EMBL/GenBank/DBJ databases.
DR Submitted (NOV-2000)
DR InterPro: IPR001103; Andrgn_receptor.
DR Pfam: PF02166; Androgen_recip; 1.
DR Receptor.
KW NON_TER
FT SEQUENCE 531 AA; 54419 MW; FA0D3AA60AC20DAB CRC64;

Query Match 57.1%; Score 2802.5; DB 4; Length 531;
Best Local Similarity 98.0%; Pred. No. 9.5e-186;
Matches 531; Conservative 0; Mismatches 0; Indels 11; Gaps 2;

Qy 1 MEVQLGLGRVYPPPSKTYRGAFQNLFSQVREVIONPGRHPEAASAAPGASLLLLQOQ 60
Db 1 MEVQLGLGRVYPPPSKTYRGAFQNLFSQVREVIONPGRHPEAASAAPGASLLLL --- 57
Qy 61 QQQQQQQQQQQQQQQQQQQQQTSRQQQQQQGGDGSQAHRRGPTGYVLVDEEQPSQ 120
Db 58 -----QQQQQQQQQQQQQQQQTSRQQQQQQGGDGSQAHRRGPTGYVLVDEEQPSQ 110
Qy 121 POSALECHPERGCVPEPAAVAAKGLPOLPAPPPEDDSDAAPSTLSLLGPTFFGLSSCS 180
Db 111 POSALECHPERGCVPEPAAVAAKGLPOLPAPPPEDDSDAAPSTLSLLGPTFFGLSSCS 170
Qy 181 ADLKDILSEASTMOLLQOQQQEAQVSSGSSGRAREASGAPTSSKDNKYLGGTSTISDNAKE 240
Db 171 ADLKDILSEASTMOLLQOQQQEAQVSSGSSGRAREASGAPTSSKDNKYLGGTSTISDNAKE 230
Qy 241 LKAVSYNGLGVEALEHLSPEQLRGDCMYAPLGVPPAVRPTPCAPLAECKGSLDDDS 300
Db 231 LKAVSYNGLGVEALEHLSPEQLRGDCMYAPLGVPPAVRPTPCAPLAECKGSLDDDS 290
Qy 301 AGKSTEDTAEYSPFKGGYTKLEGESLGCSSAAGSGSTLPLSTLSLYKSGALDEAAA 360
Db 291 AGKSTEDTAEYSPFKGGYTKLEGESLGCSSAAGSGSTLPLSTLSLYKSGALDEAAA 350
Qy 361 YQSRDYNNFLALAGPPPPPPPPHPPHARIKLENPLDYSWAAAAAQAQRYGDLASLHGAG 420
Db 351 YQSRDYNNFLALAGPPPPPPPPHPPHARIKLENPLDYSWAAAAAQAQRYGDLASLHGAG 410
Qy 421 AAGPGSGSAAASSWHTLFTAEGLYGPCGGGGGGGGGGGGGGGGGGGGEAGA 480
Db 411 AAGPGSGSAAASSWHTLFTAEGLYGPC-CGGGGGGGGGGGGGGGGGGGGEAGA 469
```

```
Qy 481 VAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPTCVKSEMGPMWDSYSGPYGD 540
Db 470 VAPYGYTRPPQGLAGQESDFTAPDVWYPGGMVSRVPYPTCVKSEMGPMWDSYSGPYGD 529
Qy 541 MR 542
Db 530 MR 531

RESULT 9
ID P70048 PRELIMINARY; PRT; 790 AA.
AC P70048;
DT 01-FEB-1997 (T-EMBLrel. 02, Created)
DT 01-JAN-1999 (T-EMBLrel. 09, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Androgen receptor alpha isoform.
GN XL ALPHA AR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Fischer L., Catz D., Kelley D.;
RT "An androgen receptor mRNA isoform associated with hormone-induced
RT cell proliferation."
RL Proc. Natl. Acad. Sci. U.S.A. 90:8254-8258(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Fischer L.M., Catz D., Kelley D.B.;
RT "Androgen-directed development of the Xenopus laevis larynx: control
RT of androgen receptor expression and tissue differentiation."
RL Dev. Biol. 170:115-126(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA Kelley D.B., Kamenetz F.R., Kelley D.B., Badea T.C.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: U67129; AAC97386.1; -.
DR HSP; P06536; IGDC.
DR InterPro: IPR001103; Andrgn_receptor.
DR InterPro: IPR000536; Hormone_rec_lig.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF02166; Androgen_recip; 1.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR000035; Znf_C4steroid; 1.
DR PRODOM: PD000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOLI; 1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 790 AA; 86973 MW; 9E7FC136CCCC3906 CRC64;

Query Match 49.2%; Score 2414.5; DB 13; Length 790;
Best Local Similarity 54.5%; Pred. No. 1.2e-158;
Matches 523; Conservative 65; Mismatches 162; Indels 209; Gaps 29;

Qy 1 MEVQLGLGRVYPPPSKTYRGAFQNLFSQVREVIONPGRHPEAASAAPGASLLLLQOQ 60
Db 1 MEVHIGLGGYKQPPGKMWIRGAFENLFLSREALQ--GERALEGSOAGWS----- 51
Qy 61 QQQQQQQQQQQQQQQQQQQQQTSRQQQQQQGGDGSQAHRRGPTGYVLVDEEQPSQ 120
Db 52 -----EAPGTHRS-----EASPDGT 68
Qy 121 POSALECHPERGCVPEPAAVAAKGLPOLPAPPPEDDSDAAPSTLSLLGPTFFGLSSCS 180
```

```

Db 69 PLNFWTHP-----PAPWRE-AQAEAPQN-PACRTE-----GAQFPALGDCP 109
QY 181 ADLKDLSEASTMQLLQQQQQAEVSGSSGRAREASGAPTSSKQNYLGGTSTISDNAKE 240
Db 110 TEUKEL-----GEQSGRIESEETP-AEKEGFSGPPEGISDSAKE 149
QY 241 LKRAVSVMGLGVALEHLSPG---BOLRGDCMYAPLLGVPVAVRTPCAPLAECCKGSL 297
Db 150 LKRAVSVMGLSMEALHLSAGAGEAQQRGDCMYAH---PPDTH--KQVAEEDKSDTR 203
QY 298 DD-----SAGKSTEDTAYSPFKGTYTKLEGESLGCSSAAAGSSGTLELPSH- 346
Db 204 DGFRRSSQSNFATGKSPED-----GGG-----GGGSSSAGSSEKQPCPD 246
QY 347 -----LSLXSGA-LDRAAAYQSDYNYFNFLALAGPPPPPPPPHP 385
Db 247 LALPEPAGGYRRAMELTSLYKTAPEESPGYPSKDFYSQMALA-----P 296
QY 386 HARIKLENLDY-GSMWAAAAQCRYGDLASLHGAAGPGSGSPSAAASSSWHTLETA 444
Db 297 HGRIKVENMEYGGGAWGAG---RYSEL-----SGFAHCGATAGWHTLF--E 339
QY 445 EGOLYPCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 504
Db 340 EQO-----SSGSFAEAG---PYSYPR-SHGPAAGADGEFPS-D 371
QY 505 VWPYG-GMYSRVPYPTSCVKSEMPWMDSYSGPYGDMRLTARDHVLPIDYPPPOKTC 563
Db 372 AWYAPTMIGRVPYSGP--MKTEMAPWMEGYPGAFGEMLRGGDRHLLPIDYPPPOKTC 429
QY 564 LTCGEASCHYGALTCGCKVFFRAAEGKQKYLCAASNDCTIDKFRKKNPCSLRKRK 623
Db 430 LTCGEASCHYGALTCGCKVFFRAAEGKQKYLCAASNDCTIDKFRKKNPCSLRKRK 489
QY 624 YEAGMTLGARKLKLGNLKLQEGEASSTSP-TEETTKLTVSHIEGYECQIFLNVL 682
Db 490 YEAGMTLGARKLKLGNLKLQEGEASSTSP-TEETTKLTVSHIEGYECQIFLNVL 549
QY 683 ATEPGVWAGHNNOPDSFALLSSNLGELQVHVYVVKAKALPGFRLNHVDDQMAV 742
Db 550 ATEPVVWAGHNNOPDSFALLSSNLGELQVHVYVVKAKALPGFRLNHVDDQMAV 609
QY 743 YSMGLMVFAMGWSFTNVNSRMFLYFAPDLVFNEMRHKSMYSQCYVMRHLHSEFGWL 802
Db 610 YSMGLMVFAMGWSFTNVNSRMFLYFAPDLVFNEMRHKSMYSQCYVMRHLHSEFGWL 669
QY 803 TTPQEFCKMALLFSIIPVDGLKNQKFFDELRMNYIKELDRYIACKRKNPTSCSRFY 862
Db 670 TTPQEFCKMALLFSIIPVDGLKNQKFFDELRMNYIKELDRYIACKRKNPTSCSRFY 729
QY 863 LTKLDSVOPIARELHQFTFDLLKSHMVSVDFFPEMMAEIIISVQPKILSGVKPIYFH 921
Db 730 LTKLDSVOPIARELHQFTFDLLKSHMVSVDFFPEMMAEIIISVQPKILSGVKPIYFH 788

RESULT 10
Q91445 PRELIMINARY; PRT; 344 AA.
ID Q91445
AC Q91445;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Androgen receptor (Fragment).
GN AR.
OS Serinus canaria (Canary).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;
OC Fringillidae; Carduelinae; Serinus.
OX NCBI_TaxID=9135;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;

```

```

RX MEDLINE=94130808; PubMed=8299561;
RA Nastluk K.L., Clayton D.F.;
RT "Seasonal and tissue-specific regulation of canary androgen receptor
messenger ribonucleic acid.";
RL Endocrinology 134:640-649(1994).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; L25901; AAA17402.1;
DR HSSP; P06536; 1GLU.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001723; Steroid_hrm_rec.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf_C4; 1.
DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
FT NON_TER 344
FT NON_TER 344
SQ SEQUENCE 344 AA; 39376 MW; 713676394FC0B030 CRC64;

Query Match 34.1%; Score 1676; DB 13; Length 344;
Best Local Similarity 90.4%; Pred. No. 4.9e-108;
Matches 311; Conservative 20; Mismatches 13; Indels 0; Gaps 0;

QY 569 EASCHYGALTCGCKVFFRAAEGKQKYLCAASNDCTIDKFRKKNPCSLRKRKYEAGM 628
Db 1 EASCHYGALTCGCKVFFRAAEGKQKYLCAASNDCTIDKFRKKNPCSLRKRKYEAGM 60
QY 629 TLGARKLKLGNLKLQEGEASSTSP-TEETTKLTVSHIEGYECQIFLNVLAEI 688
Db 61 TLGARKLKLGNLKLQEGEASSTSP-TEETTKLTVSHIEGYECQIFLNVLAEI 688
QY 689 VCAGHNNOPDSFALLSSNLGELQVHVYVVKAKALPGFRLNHVDDQMAV 748
Db 121 VCAGHNNOPDSFALLSSNLGELQVHVYVVKAKALPGFRLNHVDDQMAV 180
QY 749 MYFAMGWSFTNVNSRMFLYFAPDLVFNEMRHKSMYSQCYVMRHLHSEFGWL 808
Db 181 MYFAMGWSFTNVNSRMFLYFAPDLVFNEMRHKSMYSQCYVMRHLHSEFGWL 240
QY 809 LCMKALLFSIIPVDGLKNQKFFDELRMNYIKELDRYIACKRKNPTSCSRFY 868
Db 241 LCMKALLFSIIPVDGLKNQKFFDELRMNYIKELDRYIACKRKNPTSCSRFY 300
QY 869 SVQPIARELHQFTFDLLKSHMVSVDFFPEMMAEIIISVQPKILS 912
Db 301 SVQPIARELHQFTFDLLKSHMVSVDFFPEMMAEIIISVQPKILS 344

RESULT 11
Q93245 PRELIMINARY; PRT; 853 AA.
ID Q93245
AC Q93245;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Androgen receptor beta.
GN AR-BETA.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99150354; PubMed=10026186;
RA Takeo J., Yamashita S.;
RT "Two distinct isoforms of cDNA encoding rainbow trout androgen

```



RT J. Biol. Chem. 274:5674-5680(1999).  
RL - SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC - SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
DR EMBL: AB012096; BAA32785.1; -  
DR HSSP: P06536; 1GDC  
DR InterPro: IPR001103; Andrgn\_receptor.  
DR InterPro: IPR000536; Hormone\_rec\_lig.  
DR InterPro: IPR001628; Znf\_C4steroid.  
DR Pfam: PF02166; Androgen\_recep; 1.  
DR Pfam: PF00104; hormone\_rec; 1.  
DR Pfam: PF00105; zf-C4; 1.  
DR PRINTS: PR00047; STROIDFINGER.  
DR ProDom: PD000035; Znf\_C4steroid; 1.  
DR SMART: SM00430; HNF; 1.  
DR SMART: SM00399; Znf\_C4; 1.  
DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
KW Zinc-finger.  
SQ SEQUENCE 853 AA; 95776 MW; 65EFF5FD3B36F4C4 CRC64;  
  
Query Match 33.4%; Score 1639.5; DB 13; Length 853;  
Best Local Similarity 38.4%; Pred. No. 5.6e-105;  
Matches 386; Conservative 129; Mismatches 252; Indels 239; Gaps 30;  
  
QY 1 MEVQLGIVYPRPPKTYRGAFONLFQSVREVIONPGPRHPEAASAAPPGASLILLQQQ 60  
Db 1 MEIPVGLGVCD-SPNIVFRGPFQNVFHNKATL-----PSNTIVTETLDFSSSYFQMKN 55  
  
QY 61 QQQQQQQQQQQQQQQQQQQQQQQTSRQQQQQQQ-----EDGSPQAH-----RRGPTG 108  
Db 56 HPWEMRQNSPRKEISSTARNSDIEVKDDSDISFRTLESDDARRIHFASKSTGNKTG 115  
  
QY 109 YLVLDREQP-----SQPSALECHPERGCVPEPCAASAAKGLPQQLPAPPPDEDDSA 162  
Db 116 FSSVNELDYPNANGYSGRGPPLACNTKQCCQ-----PAPPHGVLS 159  
  
QY 163 PSTLSLLGPTFPGSLGSCADLKDILSEASTMQLLQOOQOQEAIVSEGSSGRAREAGAPTS 222  
Db 160 PNSYARVA-----NSCT----- 176  
  
QY 223 SKNYLGGTSTISDAKELCAKVSVMGLGVE--ALEHLSF-----GEOLRGDCMYAP 273  
Db 177 -----TISTARELCNAYSVSLGLTMDLNEWDGPNVAPSDANDQSGNLFQ- 225  
  
QY 274 LLGVPPAVRTPCAPLAECGSLDSDAGKSTEDTAETSPFKGYTKGLESGSCGSA 333  
Db 226 -----VPLNCSGA-----EENVSIYEK-----CPSER 249  
  
QY 334 AAGSSGTLELPSTLSLYKSGALDE-----AAVOSRDYVNFPLALAGPPPPPP 382  
Db 250 ---NARPLQSDTRVKMFKSSPANDLTBEVATMEHLSSRHPSTGEQEFRLNEKSDPTSK- 305  
  
QY 383 PHPHARIKLEPNLDYGSAAWAAACRYGLASLHGA--GAAGPG----- 425  
Db 306 -----ETENSL-----STRASASCHFDPLLAHLAHFSQTPDRISSHVIAHVCETG 354  
  
QY 426 -----SGSPSAAASSWHTLFTAEQGLYPCGGGGGGGG 461  
Db 355 ETMEDKYADYLLQOQYSVKIKYEAISNEPAGTSMGSGYNGYNDNDNTQYGP----- 404  
  
QY 462 GGGGGGGGGGGGGEAGAVA-PGYTRPPOGLAGQESDFTAPDVWYPGGMVSRVYPS 520  
Db 405 ----RQGMNPSAGPDSGFCINPYEYER--GGGLVRRR--PTSEQWYPPGMUGRMPIFNS 457  
  
QY 521 TCVKSEMPWMDSYSGPYGDMRLTARDHVLPIIDYFPPOKTCILICGDEASGCHYGALTC 580  
Db 458 PYLKNVEGDWLDV---SYTDARFEGGRDHMPMEFFPPORTCLICADEASGCHYGALTC 514  
  
QY 581 GSKVFFKRAAGKQKYLCAARNCTIDKFRKNCPCRLKCKYEAGMTLGARKLKLGN 640  
Db 515 GSKVFFKRAAGKQKYLCAARNCTIDKLRKNCPCRLKCKFEAGMTLGARKLKKIGQ 574

QY 641 LKQERGEASSTTSPTETTKL-TVSHIEG--YECQIFLNVLEAIEPGVVCAGHDNQ 697  
Db 575 LKSPPE-----DLPTQGPDTAIOCVSPQSLTFHSQLVFNILEIEPEVYNAGHDHCQ 628  
  
QY 698 PDSFAALLSLNELGERQLVYVYKAKALPGFRNLHVDDQMAVIOYSWMGLMVFAMGWS 757  
Db 629 PDSAAVLLTSLNELGERQLVYVYKAGMPGFRNLHVDDQMTVIQHSWGVVWVGLGWS 688  
  
QY 758 FTWNSRMLYFAPDLVFNEXYMRHKSRYMSQCVRMRLHSQFGLWLOITPOEFLCMKALLF 817  
Db 689 YKVNARMLYFAPDLVFNDRHMHSSMFHCIRMRHLSQOQVLLQVTOEFLCMKALLF 748  
  
QY 818 SIIPVDGLKNOKFDELURMYIKELDRIIAACKRNPTSCSRREFYOLTKLLDSVOPTAREL 877  
Db 749 SIIPVDGLKSKYFDELRLTYIINELDRVINYGRK--SNCQRFYQLTRLMDLSQPIVRKL 806  
  
QY 878 HQFTFDLLIKSHMV--SVDPPEMMAETISVQVPKILSGKVKPIYFH 921  
Db 807 QOFTFDLFIQAQSLPTKVSFPEMIAELISVHVPKILAGLAKPILFH 852  
  
RESULT 12  
Q8QFV2 PRELIMINARY; PRT; 839 AA.  
AC Q8QFV2; 80FV2;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Androgen receptor.  
GN AR  
OC Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Betka M., Rothberg S.C., Callard G.V.;  
RT "Carassius auratus Androgen Receptor";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
KW EMBL: AY090897; AAM09278.1;  
KW Receptor.  
SQ SEQUENCE 839 AA; 93168 MW; A534DC3169C0B1F4 CRC64;  
  
Query Match 33.2%; Score 1632.5; DB 13; Length 839;  
Best Local Similarity 43.9%; Pred. No. 1.7e-104;  
Matches 384; Conservative 103; Mismatches 225; Indels 163; Gaps 26;  
  
QY 130 ERGCVPEPAA-----VAASKGLPQQLPAPPDEDDSAAPSTLSLLGPTFGLSCSADL 183  
Db 44 KKGLOETDSTKRPKSLSPSKI---ICCPKKECESASSXMRSSIGSKSDTSLSCGRA 100  
  
QY 184 KDILSEASTMQLLO-----QQQOEAIVSEGSSGRAREASGAPTSSKNDYLGSTISDNAK 239  
Db 101 DATSSGSRAGLRGAESGOKSCATAEVHSRRDARVASSSRACNTTTSSSSSSSISETAR 160  
  
QY 240 ELCKAVSVSMGLGVLEALE-----HLSPGOLRDCMYAPLLGVPPAVRTPCAPLAEC 292  
Db 161 ELCKAVSVSLGAMSESLGDVEPHVPPPLTKSSERIYLFM-----PILNC 210  
  
QY 293 KGLSLDSDAGKSTEDTA-----EYSPFKGYTKGLEGE--SLGCSGA- 333  
Db 211 --SVSRXAGGKEREYALAAAGRDRAELRDRLLLEMFKSGDLEQIAGEVTLQCSSASR 268  
  
QY 334 ---AAGSSGTLELPST-----LSLYKSGA---LDEAAVQSRDYNNFLALA----- 374  
Db 269 SHLTADAQGVHEFASVSGDIANLSSEGTTPDMDTEFRASCOFEQLLPVSMAHFYQPELE 328  
  
QY 375 -GPPPPPPPPHP-----HARIKLE-NPLDYGSAW---AAACCC--RYG 411  
Db 329 NGPNQSFAPKPAEMSGFEAGFVEDYVNYLVKIAEMMPRELNDTWAYPHRYAEDCNGQYG 388



```
QY 412 DLASLHAGACGPGSGSPSARASSWHLFTTBEGQLYPCGGGGGGGGGGGGGGGGG 471
Db 389 -----SPKQRTPYASHDT-HFICNPYE----- 410
QY 472 GGGGGAAGAVPYGYGTPOGLAGQESDFTAPDVWYPGWYSRVPYPTSCVKSEMGPWM 531
Db 411 ---YGRNEALYP--RERP-----PEQWYPGMLTRPPYPMPCLKEMGNWL 453
QY 532 DSYSPYGDMLRETARHVLPIDYFPQKTCICGDRASGCHYGALTCGSKCVFFKRAA 591
Db 454 DVTS--LTDGRFDGRSDIFPMEFFLPQRTCLICSDEASGCHYGALTCGSKCVFFKRAA 511
QY 592 EGKQYLICASRNDCTIDFRKNPCSLRKCYEAGMTLGARKLKLGNLKLQEBGEASS 651
Db 512 EGKQYLICASRNDCTIDFRKNPCSLRKCFEAGMTLGARKLKLQEBGEASS 566
QY 652 TTSPT-----ETQKLTSHIEGYEQPIFLNVLAEIPGVVACAGHDNNQDPSFALLSSL 708
Db 567 -VGPVQSGSETIOCLSPKPSLTFHSQILFLNILESIEPEVVNAGHDHAQDPSAVALLTSL 625
QY 709 NELGERQLVHVVKAKALPGFRLNHLVDQMAVIOYSWMGLMVFAMGWSRFTNVNSRMLYF 768
Db 626 NELGERQLVHVVKAKALPGFRLNHLVDQMTVIQHTWGMVVFALGWSRYKNANARMLYF 685
QY 769 APDLVFNERYMHKSRMYSOCVYMRHLSOEFGLNQLITPOEFLCMKALLFSIIPVDGLKNQ 828
Db 686 APDLVFNDRHMHISMYEHCVMQKLSOEFVLLQVTOEFLCMKALLFSIIPVEGLKSQ 745
QY 829 KFFDELNMNLIKELDRITACKRNKPTSCSRFFYOLTLLDSVQPIARELHOFTEFLLIKS 888
Db 746 KYFDDLRLTYINELDLRLNLYGRK--TNCAMRFQQLTRMLDSIQPIVRKHLHOFTEFLVQA 803
QY 889 HMV--SVDFPEMMABIIISVQVPKILSGKVKPIYFH 921
Db 804 RSLPKTSFPEMABIIISVQVPKILAGLSKPILFH 838

RESULT 13
O97684 PRELIMINARY; PRT; 303 AA.
ID O97684
AC O97684
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Androgen receptor (fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DORSET;
RA Thonney M.L., Ignatz G., Richards H.M.;
RT "Partial sequence of an ovine androgen receptor cDNA.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR HSPB; AF105713; AAC97958.1; -.
DR HSSP; P06536; 1GLU.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001723; Strdhmn_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
FT NON_TER 1 1
```

```
FT NON_TER 303 303
SQ SEQUENCE 303 AA; 34673 MW; D14D3950C7DC69D1 CRC64;

Query Match 32.7%; Score 1604; DB 6; Length 303;
Best Local Similarity 99.0%; Pred. NO. 4e-103; 3; Indels 0; Gaps 0;
Matches 300; Conservative 0; Mismatches 3;

QY 564 LICGDEASGCHYGALTCGSKCVFFKRAAGCKQKYLICASRNDCTIDFRKNPCSLRKRC 623
Db 1 LICGDEASGCHYGALTCGSKCVFFKRAAGCKQKYLICASRNDCTIDFRKNPCSLRKRC 60
QY 624 YEAGMTLGARKLKLGNLKLQEBGEASSTTSPTTEETQKLTSHIEGYEQPIFLANVLEA 683
Db 61 YEAGMTLGARKLKLGNLKLQEBGEASSTTSPTTEPAQKLTSHIEGYEQPIFLANVLEA 120
QY 684 IEPGVVACAGHDNNQDPSFALLSSLNELGERQLVHVVKAKALPGFRLNHLVDQMAVIOY 743
Db 121 IEPGVVACAGHDNNQDPSFALLSSLNELGERQLVHVVKAKALPGFRLNHLVDQMAVIOY 180
QY 744 SWMGLMVFAMGWSRFTNVNSRMLYFAPDLVFNERYMHKSRMYSOCVYMRHLSOEFGLWLOI 803
Db 181 SWMGLMVFAMGWSRFTNVNSRMLYFAPDLVFNERYMHKSRMYSOCVYMRHLSOEFGLWLOI 240
QY 804 TPOEFLCMKALLFSIIPVDGLKNQKFFDELNMNLIKELDRITACKRNKPTSCSRFFYOL 863
Db 241 TPOEFLCMKALLFSIIPVDGLKNQKFFDELNMNLIKELDRITACKRNKPTSCSRFFYOL 300
QY 864 TKL 866
Db 301 TKL 303

RESULT 14
O93497 PRELIMINARY; PRT; 769 AA.
ID O93497
AC O93497
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Androgen receptor.
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Percoidae;
OC Sparidae; Pagrus.
OX NCBI_TaxID=143350;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99326178; PubMed=10395960;
RA Tsuchida K., Kinoshita M., Tokuda Y., Toyohara H., Sakaguchi M.,
RA Yokoyama Y., Yamashita S.;
RT "Sequence and expression of a cDNA encoding the red seabream androgen
RT receptor.";
RL Biochim. Biophys. Acta 1450:481-485(1999).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; AB017158; BAA33451.1; -.
DR HSSP; P06536; 1RGD.
DR InterPro; IPR001103; Andrgn_receptor.
DR InterPro; IPR000536; Hormone_rec_lig.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF02166; Androgen_recep; 1.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 769 AA; 86081 MW; 9305C627DBE42062 CRC64;
```

```

Query Match      32.7%; Score 1604; DB 13; Length 769;
Best Local Similarity 43.9%; Pred. No. 1.4e-102;
Matches 363; Conservative 79; Mismatches 171; Indels 214; Gaps 23;

QY 205 SEGSSSGRAREASGA-PTSSKDYLL----- 228
DB 46 STNGAGRMREADNADNPYTESGHMPLVCDMEKHCCQTAAAPQELFNADCRVGDERSRF 105
QY 229 GGTSTISDNALCKELKAVSYSMGLGVE-----ALEHLSPEQLRGDCMYAPLLGVPP 279
DB 106 SACATISARELCKAVSYSLGLAMESNDPSDMAALSQAANDQLRGYLFQ----- 158
QY 280 AVRPTCAPLAECKGLLDDSGAKSTEDTAEYSPKGGYTKLEGSILGSCSAAAGSSG 339
DB 159 ----VGAAPLS-CPGA-----QAAYSEYKCPPEERLHGK----- 188
QY 340 TLELPSTLSLYKSGALDEAAAYQSR-----DYNFPLALAGPPPPPPHARIKL 391
DB 189 --QQOQLMDFKS---SETGAHLQHLSTPTVDEHNFILCKAEOLTPETAHQD----- 238
QY 392 ENPLDYGSAAWAAAAOCRYGDLASLHGAAGPGS-----GSPSAAASSSWHLFTAEBOQ 447
DB 239 -----SVRAACAPY-----AQALPGNMAHFGSP--APERPW-----Q 269
QY 448 LYGPCGGGGGGG-----GGGGGGGGG----- 469
DB 270 LYKPPDEAGDFGVMESRPVTSYQPEYSVKICBEDTESALMGWNTTFNDRYNSQW 329
QY 470 -----GGGGGGGAGAVAPYGYTRPPQGLAGQESDFTAPDVMYRGMSRVYPSPCTCV 523
DB 330 GPRQCNMAHSTGANSALCHPY-----ERSVARPEHWPGML-RSPYFNSSVY 376
QY 524 KSEMGPMDSYSGPYGDMRLTARDHVLIDYFFPQKTYCLICGDBASCHYGALTCGSC 583
DB 377 KSEVGWLDV---PYSDPRFDSSEHFMPEFFPFAQRMCILCSDBASCHYGALTCGSC 433
QY 584 KVFFKRAAEKQKYLKASRNDCTIDFRKNCPCSLRKYCAGMTLGARKLKLGNLK- 642
DB 434 KVFFKRAAEKQKYLKASRNDCTIDFRKNCPCSLRKYCAGMTLGARKLKLGNLK- 493
QY 643 -----LQEEGEASSTSPTEETQKLTSHIEGYECQPIFLNLEAIPGVVYAGHDNN 696
DB 494 SDEHPLQEPAEVMPNTPSKGLS-----FNSQVFLNLESIEPEVNVAGHDY 543
QY 697 QPSFRAALLSINELGERQLVHVKKAKALPGFRLNLDVDDQMAVIOYSWGLMVFAMGR 756
DB 544 QPSAATFLTSLNELGERQLVHVKKAKALPGFRLNLDVDDQMTVIOHSMWVYVFLGWR 603
QY 757 SFTNVSRLMYFAPDLVFNRYMRHGRMYSQCYMRHLSOEFGWLQITPOEFLCKMALL 816
DB 604 SYKNVGRMLYFAPDLVFNRYMRHGRMYSQCYMRHLSOEFGWLQITPOEFLCKMALL 663
QY 817 FSIIPVDLKNQKFFDELRLNMYIKELDRITIAKRNKPTSCSRFYOLTLLDVSQPIARE 876
DB 664 FSIIPVGLKSQYFDELRLTYINELRLTYINRMN--TNCQRFYOLTRLLDVSQMTVK 721
QY 877 LQOFTFDLLIKSHW--SVDFPPEMAEILISVQPKTILSGKVRPIYH 921
DB 722 LQOFTFDLVQAOLSLPTKVSFPPEMIEILISVHVPIKILAGLAKPILFH 768

```

```

RESULT 15
Q9PWG5 PRELIMINARY; PRT: 797 AA.
AC Q9PWG5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Androgen receptor-beta.
GN AK-BETA.
OS Anguilla japonica (Japanese eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;

```

```

OC Anguilla.
OX NCBI_TaxID=7937;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=99395076; PubMed=10464240;
RA Ikeuchi T., Todo T., Kobayashi T., Nagahama Y.;
RT "cDNA cloning of a novel androgen receptor subtype.";
RL J. Biol. Chem. 274:25205-25205(1999).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AB025361; BAA83805.1;
DR HSSP: P06536; IGDC.
DR InterPro: IPR001103; Andrgn_receptor.
DR InterPro: IPR000536; Hormone_rec_lig.
DR InterPro: IPR001723; Stdhmn_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF001166; Androgen_recep; 1.
DR Pfam: PF001104; hormone_rec; 1.
DR Pfam: PF001105; zf-C4; 1.
DR PRINTS: PR00398; STRDHORMONER.
DR PRINTS: PR00047; STRODIFINGER.
DR PRODOM: PD000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOLI; 1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 797 AA; 89924 MW; CF1D5F21EEEA5691 CRC64;

Query Match      32.6%; Score 1600; DB 13; Length 797;
Best Local Similarity 43.0%; Pred. No. 2.8e-102;
Matches 379; Conservative 105; Mismatches 220; Indels 178; Gaps 27;

QY 112 LDEBOQ-----PSQPOS-ALECHPERGCVPEPGAAVAASKGLPOOLPAPDEDDSAAP 163
DB 21 LDTELQEADSANFPMQPEAEARRIHFTKS-----SDSKGDSSTL-----EPDNAQE 66
QY 164 STLSILGPTFPGLSCSADLKILSEASTMOLLOQOQOEAVERSESSSGRAREASGAPTSS 223
DB 67 NA---YGFGLDLDHDCCTDQKSLPTAPRNRINPHSDACVNSCSEA----- 113
QY 224 KDNVLTGTSTISDNALCKELKAVSYSMGLGVEALEHLSLSP-----GEQLKGDY-APL 274
DB 114 -----CTTISEARELCKAVSYSLGLNDSNEMNEPGLHNLASSIGSLDKRKNFMEVPF 166
QY 275 LGVPAPVPTPCAPLAECKGLLDDSGAKSTEDTAEYSPKGGYTKLEGE----- 325
DB 167 LG---SSGTQESVPRAEYKCARLHD--GQLLQNDLSEGMFKTTHVOGLTNEVAPSHLSS 221
QY 326 -----SLGCGSAAAGSGTLE-----LPSTLSLYKSGALDEAAA 360
DB 222 YSNMDTQTRLNATETTSASKEQASCMTMDAARSGFCQFDOLLPTTLAQISQ--IDPLSM 279
QY 361 ----YQSRDYNFPLALAGPPPPPPHFA-----RIKLENPLDY-GSAWA 402
DB 280 GRSNFRSOFYK-TLALANEAV-----EHAEGRYVDSSIQYSPKIKTENLQNSGGSW- 331
QY 403 AAAAQCRYGDLASLHGAAGPGSGPSAAASSSWHLFTAEBOQLYGPCGGGGGGGG 462
DB 332 --DHQYRYNE-----NWHYSQYGSRHGLHPYC----- 356
QY 463 GGGGGGGGGGGGAGAVAPYGYTRPPQGLAGQESDFTAPDVMYRGMSRVYPSPCTC 522
DB 357 -----TETNNQFVVEFFDQFQ---GGLLPRBSASESWFPGMLSKMPTNLPC 403
QY 523 VKSEMGPMDSYSGPYGDMRLTARDHVLIDYFFPQKTYCLICGDBASCHYGALTCG 581
DB 404 LKTEVADWLEV---PCSDRMLEGGREHVPFMEFFFPQRTCLICADEASCHYGALTCG 460
QY 582 SKVFFKRAAEKQKYLKASRNDCTIDFRKNCPCSLRKYCAGMTLGARKLKLGNL 641
DB 461 SKVFFKRAAEKQKYLKASRNDCTIDFRKNCPCSLRKYCAGMTLGARKLKLGNL 520

```



